Systematically improving espresso: insights from mathematical modeling and experiment

Michael I. Cameron and Dechen Morisco
Frisky Goat Espresso, 171 George St., Brisbane City, QLD, 4000, AUS and
Current affiliation: ST. ALi Coffee Roasters, 12-18 Yarra Place, South Melbourne, VIC, 3205, AUS

Daniel Hofstetter
Daniel Hofstetter Performance, Laenggenstrasse 18, CH-8184 Bachenbuelach, CHE

Erol Uman
Meritics Ltd., 1 Kensworth Gate, Dunstable, LU6 3HS, UK

Justin Wilkinson
Faculty of Mathematics, University of Cambridge, Cambridge, CB3 0WA, UK

Zachary C. Kennedy
National Security Directorate, Pacific Northwest National Laboratory, Richland, WA, 99352, USA

Sean A. Fontenot
Department of Chemistry and Biochemistry, University of Oregon, Eugene, OR, 97403, USA.

William T. Lee
Department of Computer Science, University of Huddersfield, Huddersfield, HD1 3DH, UK and
MACSI, Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland

Christopher H. Hendon
Materials Science Institute and Department of Chemistry and Biochemistry,
University of Oregon, Eugene, OR, 97403, USA

Jamie M. Foster
School of Mathematics & Physics, University of Portsmouth, Portsmouth, PO1 2UP, UK

Espresso is a beverage brewed using hot, high pressure water forced through a bed of roasted coffee. Despite being one of the most widely consumed coffee formats, it is also the most susceptible to variation. We report a novel model, complimented by experiment, which is able to isolate the contributions of several brewing variables, thereby disentangling some of the sources of variation in espresso extraction. Under the key assumption of homogeneous flow through the coffee bed, a monotonic decrease in extraction yield with increasingly coarse grind settings is predicted. However, experimental measurements show a peak in the extraction yield versus grind setting relationship, with lower extraction yields at both very coarse and fine settings. This result strongly suggests that inhomogeneous flow is operative at fine grind settings, resulting in poor reproducibility and wasted raw material. With instruction from our model, we outline a procedure to eliminate these shortcomings.

Lead Contact: chendon@uoregon.edu
jamie.michael.foster@gmail.com

I. INTRODUCTION

The past century has seen an increase in the prevalence of coffee consumption, as consumers have gained an appreciation for its complex and exciting flavours, and obvious psychological effects[1–7]. As a result, the coffee industry significantly contributes to the economic stability of numerous producing and consuming countries. For example, in 2015 the American coffee industry provided over 1.5 million jobs, accounting for $225.2 billion USD (1.6% gross domestic product), and resulting in *ca.* $30 billion USD in tax revenue.[8] However, coffee producing countries now face new challenges owing to changing climate[9–12] and shifts in consumer preferences. These challenges highlight the need to maximize the quality and reproducibility of the beverage while minimizing the mass of coffee used to produce it.

Of all of the coffee formats, espresso is by far the most complicated and susceptible to fluctuations in beverage quality. As historically defined by the Specialty Coffee Association, an espresso is a 25 - 35 mL (*ca.* 20 - 30 g) beverage prepared from 7 - 9 g of ground coffee made with water heated to 92 - 95 °C, forced through the granular bed under 9 - 10 bar of static water pressure and a total flow time of 20 - 30 s. These metrics have been grandfathered into the industry, and are significantly detached from the recipes used in most cafes today. Coffee shops routinely favor higher dry coffee mass (15 - 22 g), resulting in larger volume beverages (30 - 60 g beverage mass), produced on machines that dynamically control both water pressure and temperature. The variables of tamp force, flow rate/time, dry mass of coffee, and beverage volume are all determined by the machine's operator.

There are other variables that impact the beverage quality prior to the ground coffee being exposed to water. The grind setting determines the particle size distribution of the coffee grounds (and therefore the surface area).[13] Once compacted into a granular bed, the particle size distribution plays a role in controlling the permeability of the bed and consequently the flow rate. One can achieve a decreased flow rate in a number of ways: by decreasing water pressure, grinding finer, packing the bed more tightly, using more coffee, or some combination thereof. A further source of variability is that roasted coffee ages through off-gassing, losing roast-generated volatiles thereby altering the resultant beverage density and flavor[14,15].

In principle, it is preferable to make objective statements about the flavor of foodstuffs from knowledge of their molecular components. This poses problems for coffee as there are there are ∼2000 different compounds extracted from the grounds during brewing[16,17]. In practice, we are limited to more easily measurable descriptors. The coffee industry uses extraction yield (EY), a ratio of solvated coffee mass to the mass of dry coffee used to produce the beverage, to assess extraction. EY is calculated by first measuring the refractive index, a property that depends on temperature. While a refractive index measurement cannot be used to characterize the beverage composition (*i.e.* it cannot be used to make qualitative statements about chemical composition; the refractive response is highly molecule specific)[18], it has been shown to accurately correlate with extracted mass[19]. This is turn may be related to flavor for a narrow range of brew parameters; we will discuss this further in subsequent sections. Accordingly, the Specialty Coffee Association advises that coffee most frequently tastes best when the proportion of extracted dry mass is in the range 17–23%. Coffee beverages with EYs exceeding 23% typically taste bitter, while those below 17% are often sour. Furthermore, it should be noted that concentration (often referred to as beverage strength) plays another key role in coffee beverage production. Here, we consider this a secondary problem and chose to monitor EY as it is still a descriptor of flavor, but also has significant economic implications (*i.e.* it tells us something about how efficiently we are using our coffee mass). In contrast, one could argue that beverage concentration is related to the consumers preference.

This paper reports the development of a multi-scale mathematical model for extraction from a granular bed. Here, multi-scale is used to emphasize the fact that the descriptions of the physics spans different length scales (*i.e.* the size of the coffee grain which is much smaller than the size of the espresso bed)[20]. We apply the model to espresso-style coffee extraction, but note that it is readily generalizable to any liquid/granular biphasic system. The model offers scope to independently alter familiar variables such as grind setting, water pressure, flow rate, coffee dose, extraction kinetics, and so forth: these culminate in a prediction of EY. The model's ability to individually change each brewing parameter is crucial to developing enhanced understanding of brewing because altering parameters truly independently is difficult in an experimental setting.

The model enables us to understand the origin of irreproducibility in espresso (namely non-uniform flow) and it also informs us in proposing a novel strategy for minimizing drink variation as well as dry coffee waste. We identify a critical minimum grind size that allows for homogeneous extraction. Below this setting, a counterintuitive reduction in EY and increase in variability is observed. In a departure from the SCA recommendations, the model suggests that we should ignore brew time and navigate the EY landscape using only mass of coffee and mass of water as independent variables. We demonstrate that we are able to systematically reduce coffee waste and dramatically reducing shot variation, while also saving the cafe both time and money in their production of espresso-based beverages. Our approach is then implemented into a real cafe setting where the economic benefits were monitored. Using these data paired with those previously reported,[8,21] we estimate that a 25% reduction in coffee mass per coffee beverage will result in approximately *ca.* $3.1 million USD per day.

II. DEVELOPMENT OF A RATIONAL MODEL FOR ESPRESSO EXTRACTION

Espresso is brewed in a cylindrical container denoted by $z \in (0, L)$ and $R \in (0, R_0)$, Figure 1. The solid coffee grounds occupy part of the cylinder, Ω_s, and contain a concentration of soluble coffee, c_s. The cylinder also contains inter-granular pore space, Ω_l, which is occupied by liquid during extraction, which itself contains a concentration of coffee solubles c_l. We use the term coffee solubles to denote the sum of the concentrations of all compounds in coffee; this is in line with the EY measurement. We note that the model could readily be generalized to explicitly track any number of chemicals. However, the utility of doing so is questionable because one would need to also provide/fit kinetic parameters for each individual compound, rendering the model susceptible to over-fitting. We would also require knowledge of initial concentrations and these are difficult to measure for many species. The liquid flow between the inlet ($z = 0$) and outlet ($z = L$) is driven by an over-pressure (the pressure excess relative to atmospheric pressure, P_{tot}), applied by a pump. The model equations take the form of a system of partial differential equations that describe, i) the transport of coffee solubles from the interior of the grounds to their surface, ii) the exchange/dissolution of the solubles from the grounds into the liquid, and iii) the migration of the solubles in the liquid by a combination of diffusion and advection.

The solubles in the liquid phase are transported by a combination of diffusion and convection due to the flow of the liquid through the bed. The concentration of solvated coffee is therefore governed by an advection-diffusion equation

$$\frac{\partial c_l}{\partial t} = \nabla \cdot (D_l \nabla c_l - \mathbf{u} c_l) \quad \text{in} \quad \Omega_l, \tag{1}$$

where t, D_l and $\mathbf{u}$ are time, the diffusivity of solubles within liquid, and the velocity of the liquid respectively. The liquid flow is solved for via the Navier-Stokes equations

$$\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u} = -\frac{1}{\rho}\nabla P + \frac{\mu}{\rho}\nabla^2 \mathbf{u}, \tag{2}$$

$$\nabla \cdot \mathbf{u} = 0 \quad \text{in} \quad \Omega_l, \tag{3}$$

where μ, ρ and P are liquid viscosity, density, and over-pressure respectively.

Work by Spiro and colleagues demonstrated that the transport of coffee solubles through the interior of the grounds can be described by a diffusion process.[22–28] Hence

$$\frac{\partial c_s}{\partial t} = \nabla \cdot (D_s \nabla c_s) \quad \text{in} \quad \Omega_s, \tag{4}$$

where D_s is the diffusivity of solubles within the grains. Here, we treat coffee grounds as spherical dense particles, but we note that the coffee grains themselves may be irregularly shaped and feature intragranular macropores, as previously observed in scanning electron micrographs[29]. As discussed in the coming section, the majority of particles in ground coffee are smaller the macropore diameter observed the micrographs. Nitrogen physisorption was used to assess the microporosity of the coffee grounds; the data suggests that ground coffee does not feature microporosity (see Supplemental Information). Thus we expect our description to hold for most espresso grind settings.

Boundary conditions at the inlet, $z = 0$, include a specified fluid over-pressure, the requirement that the water enters the basket with a purely normal velocity and that the normal flux of dissolved species should be zero

$$P|_{z=0} = P_{tot}, \tag{5}$$

$$\mathbf{u} \cdot \hat{\mathbf{t}}|_{z=0} = 0, \tag{6}$$

$$(-D_l \nabla c_l + \mathbf{u} c_l) \cdot \hat{\mathbf{n}}\Big|_{z=0} = 0, \tag{7}$$

where $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$ are the unit vectors tangent and normal to the surface $z = 0$ respectively. At the exit we apply conditions of zero over-pressure, zero tangential velocity and a condition that there is zero diffusive flux of coffee. In summary

$$P|_{z=L} = 0, \tag{8}$$

$$\mathbf{u} \cdot \hat{\mathbf{t}}|_{z=L} = 0, \tag{9}$$

$$(-D_l \nabla c_l) \cdot \hat{\mathbf{n}}\Big|_{z=L} = 0. \tag{10}$$

At the vertical edges of the cylinder, $R = R_0$, no flux conditions are applied to the liquid coffee concentration as the liquid cannot exit in these directions

$$(-D_l \nabla c_l + \mathbf{u} c_l) \cdot \hat{\mathbf{n}}|_{R=R_0} = 0, \tag{11}$$

$$\mathbf{u}|_{R=R_0} = \mathbf{0}. \tag{12}$$

On the boundaries between the grains and inter-granular pore space, Γ_{int}, there is a flux of solubles per unit area which we denote by G. Appropriate boundary conditions are

$$(-D_s \nabla c_s) \cdot \hat{\mathbf{n}} = G, \tag{13}$$

$$(-D_l \nabla c_l + \mathbf{u} c_l) \cdot \hat{\mathbf{n}} = G, \tag{14}$$

$$\mathbf{u} = \mathbf{0} \quad \text{on} \quad \Gamma_{int} \tag{15}$$

where the former two capture mass transfer and the latter imposes that the liquid should be stationary on the grain/pore-space interface.

Determining the form of the reaction rate, G, is non-trivial and it is something that is not readily measured experimentally. However, it can reasonably assumed that the rate of transfer of solubles between the phases should depend on the local concentrations of solubles near the interface. Furthermore, the rate of extraction is zero when: i) the liquid immediately outside the grain is saturated (*i.e.* at a concentration c_{sat}) or, ii) when the liquid outside the grain is at the same concentration as the grain (*i.e.* in equilibrium) or, iii) when the grain is depleted of solubles (the experimental upper limit of extraction is

approximately 30 % by mass). We therefore postulate a rate which satisfies all of the above conditions, namely

$$G = kc_s(c_s - c_l)(c_{sat} - c_l) \quad \text{on} \quad \Gamma_{int} \tag{16}$$

where k is a rate constant. We note that the quantity c_{sat} likely depends on the local temperature. One could readily incorporate a thermal model into the description, but here we assume that the espresso basket is isothermal. This is justified on the basis that the heat capacity of water is relatively high and that espresso basket temperatures are actively controlled in most machines.

Coffee particulates remain dry until they are connected to the extraction apparatus, at which point water is rapidly introduced to the bed, serving to wet the entire puck and stabilize the particle temperature. Modeling this initial wetting (*i.e.* pre-infusion) stage poses another series of interesting problems; the model presented here is only valid once liquid infiltration has taken place and we refer the interested reader to[30] where pre-infusion has been discussed. We avoid explicitly modeling this stage by assuming that at $t = 0$, when extraction begins, the bed is filled with liquid water that is free from solubles. We therefore have

$$c_l|_{t=0} = 0, \qquad c_s|_{t=0} = c_{s0}, \tag{17}$$

and note that the errors engendered in making this approximation can be expected to be small because the intrusion stage represents only a small portion of the overall extraction time. In (17) c_{s0} is the concentration of solubles in the grains initially. Concurrent with the wetting stage is the potential for the grains in the bed to be rearranged by the invading fluid[28]. Rearrangement that may occur during the initial wetting stage will be accounted for later after the equations have been homogenized. One of the results of this procedure is that the geometry is encapsulated in the macroscopic quantity of permeability, and by making this material property inhomogeneous the model can mimic a non-uniform distributions of grounds.

A. Particle size distribution of ground coffee

The model requires knowledge of the distribution of coffee particle sizes produced by the grinder. The population, surface area, and volume fraction of the particles are used to estimate the permeability of the bed, and this is crucial in determining the liquid flow. Moreover, the particle size controls the extraction dynamics because it determines the typical distance (and in turn the typical time) over which solubles must be transported within the grains before they reach the interface where they can be dissolved into the liquid.

Particle size distributions were measured using our described experimental procedure; these data are shown in Figure 2a. We observe that, similar to the formation of two families of particles sizes found in exploding volcanic rock[31], there are two groups of particle sizes in ground coffee. Namely, boulders (which we define as larger than 100 μm) and fines (smaller than 100 μm in diameter). This bimodal distribution is caused by large particles fracturing until they are sufficiently small to exit through the grinder burr aperture.[13] The size of the boulders are determined by the burr separation themselves, while the fines (much smaller than the burr aperture) are thought to be produced at the fracture interface. One piece of evidence supporting this idea is that as the grind setting, G_S, is reduced, the relative proportion of fines increases, but their size remains constant, Figure 2b.

B. Multi-scale homogenization and one-dimensional reduction

Direct solution of (1)-(17) on a realistic packed bed geometry comprised of many millions of individual grains is intractable, even using modern high performance computing. Therefore, rather than tackling the problem directly we make use of the vast disparity in the length-scales between that of a coffee grain ($\sim$10 μm, referred to as the microscopic scale) and that of the whole bed ($\sim$1 cm, referred to as the macroscopic scale) to systematically reduce the system using the asymptotic technique of multiple scales homogenization. Such techniques have been applied to problems with a very similar structure in electochemistry[32] and rather than present this very involved calculation in full here, we provide a summary in the Supplemental Information and refer the interested reader to Richardson and co-workers[33] where the the details of an analogous calculation are presented.

The macroscopic system of equations, valid on the larger macroscopic length scale of the entire bed, systematically follow from the microscopic equations (1)-(17). The application of the multiple scales technique significantly reduces the model complexity, but the requisite information about the microscale variations is retained. For example, since the dissolution rates depend on the concentration of solubles on the microscopic particle surfaces, the multi-scale system contains a series of microscale transport problems that must be solved inside representative grains. It is crucial that this microscopic information is preserved in the multi-scale model because it will allows us to study the effects of different grind settings on the overall macroscopic behavior of the extraction.

Motivated by the bimodal distribution of particle sizes the model it may be assumed that the bed is comprised of two families of spherical particles with radii a_1 (fines) and a_2 (boulders). Further, we denote the Brunauer-Emmett-Teller (BET) surface area of the different families of particles by b_{et1} and b_{et2} respectively. The BET surface area characterizes to amount of interfacial surface area between two intermingled phases per unit volume of the mixture, and therefore has units of 1/m. We also introduce c_{s1} and c_{s2} to denote the concentrations of solubles in the two particle families. The resulting macroscopic equation for the concentration of solubles in the

liquid is

$$(1-\phi_s)\frac{\partial c_l^*}{\partial t} = \frac{\partial}{\partial z}\left(D_{eff}\frac{\partial c_l^*}{\partial z} - qc_l^*\right) + b_{et,1}G_1 + b_{et,2}G_2. \tag{18}$$

Here the quantity c_l^* is the concentration of solubles in the liquid as predicted by the multi-scale modeling approach; whereas c_l appearing in the equations in the Supplemental Information is the concentration of solubles in the liquid as predicted by the original microscopic model. The upscaled and reduced versions of (7) and (10) are

$$-D_{eff}\frac{\partial c_l^*}{\partial z} + qc_l^*\Big|_{z=0} = 0, \tag{19}$$

$$-D_{eff}\frac{\partial c_l^*}{\partial z}\Big|_{z=L} = 0. \tag{20}$$

These assert that there should be no flux of solubles across the inlet and no diffusive contribution to the flux at the outlet. In the next section we will show that parameter estimates indicate that diffusive fluxes are negligible compare to those due to advection in typical espresso brewing conditions. Hence, it is the flow of the liquid through the pores that is primarily responsible for moving solubles through the bed once they have been dissolved. Hence, even though the physical relevance of the latter condition in (19) is not completely clear, it has negligible impact on the model solution.

The microscopic equations to be solved are

$$\frac{\partial c_{si}}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_s\frac{\partial c_{si}}{\partial r}\right), \quad \text{for} \quad i = 1, 2, \tag{21}$$

and the symmetry and dissolution rate boundary conditions, which arises from (13), and which acts to couple the micro- and macroscale transport problems, are

$$-D_s\frac{\partial c_{si}}{\partial r}\Big|_{r=0} = 0, \tag{22}$$

$$-D_s\frac{\partial c_{si}}{\partial r}\Big|_{r=a_i} = G_i, \quad \text{for} \quad i = 1, 2.$$

The problem is closed by supplying the initial conditions (17) and the reaction rates, G_i. This has precisely the same form as (16) but with additional subscripts to differentiate the boulders from the fines, *i.e.*, $G_i = kc_{si}(c_{si} - c_l^*)(c_{sat} - c_l^*)$.

A formula for EY in terms of the model variables can be derived by first noting that if follows from (18) and (20) that an expression for the rate at which *soluble mass* enters the cup is given by

$$\frac{dM_{cup}}{dt} = \pi R_0^2 q c_l|_{z=L}. \tag{23}$$

On integrating this equation along with the initial condition that there is no solvated mass at $t = 0$, and dividing by the dry mass of coffee initially placed in the basket, M_{in}, gives

$$\text{Extraction yield (EY)} = \frac{\pi R_0^2 q \int_0^{t_{shot}} c_l|_{z=L} dt}{M_{in}}, \tag{24}$$

where EY is described as the fraction of solvated mass compared to the total mass of available coffee. Here, t_{shot} is the flow time. Equation (24) will be used in the coming sections as a means to compare model predictions of EY to experimental measurements.

C. Tuning the model to espresso extraction data

Initially, simulations of espresso extraction were run using a cafe-relevant recipe of 20 g of dry grounds used to produce a 40 g beverage, under 6 bar of static water pressure. Values for the radius of an espresso basket (R_0), the viscosity of heated water (μ), the saturation concentration of heated water (c_{sat}) and the concentration of solubles initially in the grounds (c_{s0}) are readily available in the literature, see the Supplemental Information for a summary of values and their sources. Moroney and colleagues[30] estimate that the volume fraction of grounds in a packed bed is $\phi_s = 0.8272$ and this, along with the density of grounds and the bed radius allow us to derive a value for the bed depth via the relationship

$$\pi R_0^2 L = \frac{M_{in}}{\rho_{grounds}}\phi_s. \tag{25}$$

Whilst it is likely that bed depth varies slightly across the range of grind settings (as the volume fraction changes), we assume that the bed depth is constant for a given dry mass of coffee, M_{in}. Values for both the radii and BET surface area for the two differently sizes families of particles in the grounds can be derived from the data shown in Figure 2 provided that both families are distributed homogeneously throughout the bed. The Darcy flux, q, determines the flow rate of the liquid through the bed and varies with grind setting. They are estimated using the shot times presented in Figure 4 using the equation

$$q = \frac{M_{out}}{\pi R_0^2 \rho_{out} t_{shot}} \tag{26}$$

where M_{out} is the mass of the beverage (40 g), and we make the assumption that the density of the drink, ρ_{out}, is the same as that of water but we note that this is an area which we could be improved in future model developments. We emphasize the difference between M_{out} and M_{cup}; the former is the total mass of the beverage, whereas M_{cup} (used in (23)) is the total mass of solubles in the beverage. The parameter values discussed above are tabulated in the tables presented in the Supplemental Information.

Three parameters, namely D_{eff}, D_s and k, remain to be specified. The effective macroscopic diffusivity in the liquid is often related to the diffusivity D_l via

$D_{eff} = \mathcal{B}D_l$ where $\mathcal{B}$ is the permeability factor. This accounts for the reduction in the diffusive fluxes due to the obstacles provided by the intermingled phase, in this case the coffee grains. This permeability factor can either be computed via a homogenization calculation[34] or can be estimated using the Bruggemann approximation which asserts that $\mathcal{B} = \epsilon_l^{3/2}$.[35] Unfortunately, neither D_l nor D_{eff} have been experimentally characterized. However, if one adopts a value for the diffusivity of a typical compound in water and then compares the expected size of the flux of solubles due to diffusion versus convection, it is seems clear that a safe conclusion is that the former is significantly smaller than the latter

$$\frac{D_{eff}c_{sat}}{L} \ll qc_{sat}. \tag{27}$$

We note that this same conclusion was also reached previously[30]. Henceforth, we assign the macroscopic diffusivity of solubles in the liquid with a small value, so that diffusive transport is negligible compared to convection due to liquid flow. The final two parameters, D_s and k shall be fitted to experiment. The results of this fitting are shown in Figure 4 and leads to values of

$$D_s = 6.25\times10^{-10}\mathrm{m}^2/\mathrm{s}, \quad k = 6\times10^{-7}\mathrm{m}^7\mathrm{kg}^{-2}\mathrm{s}^{-1}. \tag{28}$$

D. The effect of altering brew ratio and water pressure

Under the assumption that the bed geometry depends only on grind setting, it is straightforward to explore the role of altering both dry coffee mass (*i.e.*, the brew ratio) and static water pressure. When the coffee mass is altered the only parameter that needs to be altered is the bed depth, L. We continue to operate under the assumption that the bed has a fixed volume fraction of coffee grounds and hence the depth of the bed is directly proportional to the coffee dose, see (25) and the Supplemental Information. When the pump overpressure (P_{tot}) is increased, the Darcy flux (q) is increased in direct proportion,[34] whereas the shot time is decreased in an inverse proportional manner, (26) and the Supplemental Information. We monitor the extraction yield (the mass fraction of grains that are dissolved), enabling the isolation of both coffee mass and overpressure. These results are shown in Figure 3.

III. BREWING ESPRESSO

The model predicts that extraction yield can be increased by grinding finer, using lower pressure water, and/or using less coffee. The model was then compared to experimental coffee brewing, performed in a cafe setting. Using an espresso machine set to $P = 9$ bar of water pressure resulted in clogging at fine grind settings (see the Supplemental Information). To circumvent this difficulty the pressure was reduced to $P = 6$ bar. Coffee must be tamped to level the granular bed. We explored a range of tamp pressures, but did not observe an appreciable variation in shot time or extraction yield and so we standardized our tamping procedure using an automated device that pressed the bed at 98 N (see the Supplemental Information). Flavor differences, however, were noted but not quantified. The barista will need to taste their coffee in the cafe setting to ensure the beverage has the qualities that they desire. In summary, however, lower water pressure and tamp force allowed for access to a wider range of grind settings, thereby allowing for systematic sampling of shot time and extraction yield over all relevant espresso grind sizes (Figure 4). These results indicate that the relationship between shot time is linear (Figure 4a); with a coarser grind setting resulting in shorter shot times.

Examination of the extracted mass of coffee as a function of grind setting reveals a more puzzling outcome (Figure 4b). From our model it was anticipated that decreasing the grind setting should increase extracted mass because the grinder, i) produces more fines, yielding higher surface area, ii) produces smaller boulders, reducing the length of transport pathways for solubles from their interior to their surface and, iii) increases shot times and, in turn, contact time and allows more time for dissolution of coffee compounds. This counter-intuitive decrease in extraction yield with grind settings less than 1.7 indicates that regions of the bed are not being evenly extracted (*i.e.*, flow is no longer homogeneous). The extraction yield measurements are made based on a sample of the beverage from the brewed cup of coffee, and is hence indicative of an "average" extraction yield of the grains throughout the bed. Thus, the onset of nonhomogeneous flow should be accompanied by a perceived mixture of both under- and over-extracted flavors; this experience is very familiar in specialty coffee. Consumers may describe the same espresso as tasting both bitter and acidic (orthogonal flavors originating from dissimilar chemical motifs, particularly detectable as the coffee cools)[36,37]. It is also worth noting that the industry often uses grind settings less than 1.7, in part to hit the time targets described in the definition of espresso.

The non-monotonic trend in extraction yield with G_S can be attributed to two competing extraction regimes, namely i) the expected flow conditions where extraction increases as the coffee is ground finer, and ii) aggregation and/or inhomogeneous density in the bed causing partially clogged flow and reduced average extraction. This highlights a fundamental problem in correlating a coffee beverage with refractive index measurements as there are countless ways to achieve a given EY. For example, examining our data reveals that one can obtain two 22 % EY shots by keeping the brew ratio fixed, and setting the G_S to either *ca.* 1.3 or 2.0. The chemical composition of the faster shot cannot be the same as the slower shot owing to molecular differences in solubility, dissolution rate, and resultant molecule-dependent impact on

refractive index[38]. This result does not undermine our use of EY, but rather illustrates that the barista indeed needs to taste their coffee, rather than measure its solvated mass.

Our results provide an avenue to tackle three highly relevant issues in the coffee industry: first, how can one improve espresso reproducibility given the nonlinear dependence of extraction yield on grind setting? Second, what should one do if they want to reduce shot time or extraction yield variability? And third, can we systematically improve espresso reproducibility while minimizing coffee waste? We will address these questions in Section V.

IV. ACCOUNTING FOR THE PARTIALLY CLOGGED FLOW REGIME

The discrepancy between the model and experiment for finely ground coffee ($G_S < 1.7$, Figure 5), allows us to estimate the amount of the bed that is effectively inaccessible due to clogging. For an extraction in the partially clogged regime, flow is non-uniform and this causes some regions of the granular bed to be more extracted than others. Without a precise map of the bed geometry it is difficult to characterize the partially clogged flow pattern precisely. However, one can imagine an extreme case of partial clogging in which some regions of the coffee bed have zero flow (and are therefore non-extracted), while the remainder is subject to homogeneous flow. Based on a comparison between the model prediction and the experimental measurement, and assuming this extreme case of partial clogging (*i.e.*, there are regions in the coffee bed that are entirely dry), we find that there is a 13.1%, 6.1% and 2.6% difference in predicted EY versus experimental values, for grind settings of 1.1, 1.3 and 1.5, respectively. Of course, it is likely that the clogging results in a spread of extraction yields, with a portion of the mass extracted to 0%, 1%, 2% etc. Thus, estimating the level of inefficiency in a given extraction is not possible using the refractive index measurement alone, and we hope that follow-up studies will provide molecular handles on the disparities in flavor comparing homogeneous and partially clogged extractions with identical EYs.

We are able to adapt the model in order to recover this downward trend after reaching the critical grind setting. In essence, this is achieved by reducing accessible surface area of a fraction of the granular bed, simply modeled by reducing the surface area of the planar faces of the cylindrical bed that is exposed to incident water. This procedure was performed empirically until our predicted EY matched the experimental data. More important, however, are the implications of the need to reduce the proportion of coffee that is accessible in the model. First, this suggests that in many circumstances where flow is inhomogeneous there are regions of the granular bed that have been extract far higher than measured with an average EY. Second, the grind setting plays a major role in determining how much dry coffee mass is wasted in the brewing process. Finally, marriage of the model and experiment provide us one clear avenue to optimizing espresso extraction.

V. SYSTEMATICALLY IMPROVING ESPRESSO

Since the extent of the clogging is determined by the size distribution in the grind, and the variation in how different coffees grind are negligible[13], it is reasonable to expect all coffees to exhibit this inhomogeneous "peaked" relationship between extraction yield and grind setting for a fixed pump pressure and brew ratio. We therefore use the insights gleaned by comparing the model to experiment above to make generic recommendations on how to address the important economic aims of maximizing extraction yields, while simultaneously producing enjoyable espresso and reducing dry coffee waste. The following sections discuss several approaches to achieving these goals.

A. Maximizing extraction yield by altering G_S

The grind setting that gives rise to the maximum extraction yield, EY_{max}, for a given set of brewing parameters should correspond to the finest grind that maintains homogeneous extraction from the coffee bed. Since, by definition, EY_{max} is greater than or equal to the EY obtained for the extraction parameters used in cafes, the barista can always find their targeted extraction yield by first extracting at EY_{max}, then changing the brew ratio by increasing the volume of water used to produce the shot. In practice, this is achieved by first locating a so-called tasty point (*i.e.*, an espresso shot that tastes good to the barista), then locating the grind setting corresponding to EY_{max} using either the refractive technique detailed here, by simply tasting the coffee. Following Figure 6a, the operator must only adjust the grind setting to find EY_{max} (shown in green), followed by reducing the volume of water used in the extraction to result in the same EY as measured using the refractive index device (shown in red). One should note that this process results in the reduction in total cup volume, but does increase the cup concentration and reproducibility.

B. Systematic reduction of coffee mass by down-dosing and grinding coarse

As we demonstrated in Figure 3, our model informs us that a reduction in dry coffee mass results in an increased EY_{max} (shown schematically in blue in Figure 6b). Thus, a barista is able to achieve highly reproducible espresso with the same EY as the 20 g espresso, by reducing the coffee mass to 15 g, and counter-intuitively

grinding much coarser (as shown in red, Figure 6b). This modification may result in very fast shots (<15 s), a reduction in espresso concentration, and a different flavor profile.

The Specialty Coffee Association espresso parameters mandate that the extraction should take 20-30 seconds, we speculate that this might be partially responsible for the prevailing empirical truth that most coffee is brewed using grind settings that cause partially clogged/inhomogeneous flow. Remembering that the initial tasty point may lay in the clogged flow regime, some of the bed is extracted much more than the refractive index measurement suggests. By lowering the dry coffee mass and grinding to maximize EY, the operator may notice that they are able to push their extractions much higher than before, while achieving highly reproducible espresso. Indeed, the two approaches presented in Figure 6 are complimentary as the former increases shot concentration, while the latter decreases dry mass. There are circumstances where businesses make decisions such as the minimum concentration and volume of espresso that is acceptable to present to customers. When iterated, these approaches result in a optimization of beverage volume, concentration, and other economic implications.

C. Blending shots

Beyond sensory science studies, a persistent difficulty is that there is no rapid routes to assessing qualities of two identical extraction yields made with different grind settings or brew parameters. It is clear that espresso made at 22% EY in the partially clogged regime tastes more "complex" than a fast 22 % EY obtained using the optimization routine presented in Figure 6. In an attempt to recover the same flavor profile as the partially clogged flow regime a shot must contain a mixture of higher and lower extractions. Consider the tasty point in Figure 7: One can approximate it's flavor profile by blending two shots, i) a low extraction/high dose (purple point), and ii) a high extraction/low dose (green point). This procedure can more economically yield a shot with an flavor profile that should approximate that which was previously only obtainable in an economically inefficient partially clogged shot. It should be noted that blending shots does double the total volume of the beverage and the procedure comes with the added combinatorial complexity associated with calibrating two shots that, when mixed together, yield superior flavor. We expect only the most enthusiastic practitioners would consider this approach but it may well be actionable in an industrial setting where extraction is carried out in bulk.

VI. IMPLEMENTATION AND OUTLOOK

Finally, in 2017 we implemented the waste reduction protocol into a local specialty cafe in Eugene, Oregon. Examination of their sales data between September 2018 – 2019 revealed that the cafe produced 27,850 espresso-containing beverages. Previously, each beverage would have contained 20 g of coffee dry mass. This specialty grade grade coffee is valued at $0.53 per 20 g. By systematically reducing this mass by 25%, the cafe was able to save $0.13 per drink, amounting to a revenue increase of $3,620 per year. In addition to the monetary saving, the shot times were routinely reduced to 14 seconds, significantly reducing the order-to-delivery time. From this proof of concept, we can speculate on the larger economic benefit gained by the procedure detailed herein. Encompassing both specialty and commodity grade coffee beverage products we estimate that the average coffee beverage is produced using *ca.* $0.10 of coffee. Considering a 25% reduction in coffee mass (*i.e.* $0.025 saving), and considering the daily coffee consumption in the USA (124,000,000 espresso-based beverages per day)[21], our protocol yields $3.1 million USD savings per day, or $1.1 billion USD per year. Of course, this poses significant problems for the entire supply chain, as being more efficient with ground coffee does yield less revenue for roasters, importers, and producers, and further highlights that there is still much work to be done to improve efficiency in the industry, while also uniformly increasing profits.

While we do not present solutions to all of these interesting problems, we have described the formulation of a novel model for extraction (*i.e.* mass transfer) from a granular bed composed of mixed particle sizes to a liquid which flows through this bed. Multiple scales homogenization, which exploited the disparity between the length scales associated with individual grains and those of the whole bed, has been used to reduce the model on its original intricate geometry to a multi-scale model which has markedly less geometric complexity and is therefore usable and more readily diversified. The model is able to faithfully reproduce experimental measurements in regimes where flow is homogeneous and accurately predicts extraction yields as a function of meaningful parameters such as coffee and water masses, extraction rates. While the extraction yield is not directly indicative of quality, it does allow for economic arguments to be made. Furthermore, such measurements, paired with model predictions, have lead to novel insight which suggests several strategies for systematically improving espresso reproducibility as well as reducing coffee waste, leading to a more sustainable production of high quality beverages. Ultimately, we conclude by presenting a route to obtaining complex flavor profiles whilst maintaining economic savings through blending of shots. While the latter is not necessarily practical, it does highlight that partially clogged flow may impart complexity, albeit with large variation.

VII. ACKNOWLEDGMENTS

This work was enabled by charitable donations of equipment Barista Technology BV (puqpress), Acaia Corp. (balance), Frisky Goat Espresso (personnel and coffee), and our continued collaboration with Meritics Ltd. (particle size analyses). The authors are grateful to M. and L. Colonna-Dashwood, A. Thomas Murray, and R. Woodcock for insightful discussions. We thank Tailored Coffee Roasters, Eugene, OR, for their implementation of the procedures detailed herein, and B. Sung and M. Pierson providing their revenue data. We are also grateful to Death Wish Coffee for providing us with the image used in the graphical abstract. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant number ACI-1053575.

VIII. DECLARATIONS OF INTEREST

The authors declare no competing financial interests.

IX. EXPERIMENTAL PROCEDURES

Espresso was prepared using standard equipment at Frisky Goat Espresso. 20 g ridge-less baskets were fitted into the porta-filters of a San Remo Opera three group espresso machine. The Opera allows for precise control of shot time, water pressure (P_W), and temperature. Coffee was ground on a Mahlkönig EK 43 grinder fitted with coffee burrs. Espresso is typically ground at grind setting (G_S, arbitrary units) = 1.3 - 2.3, depending on the coffee. Tamp force (t_F) was controlled using the Barista Technology BV Puqpress, an automated tamper accurate to within ±3 N. We elected to use an espresso profile specialty coffee, roasted by Supreme Roasters (Brisbane, Australia). The mass of coffee in the basket and the mass of the outgoing liquid coffee were measured on an Acaia Lunar espresso balance.

Although an exhaustive characterization of the chemistry in each shot is necessary for the absolute description of shot composition (and therefore quality) we can use the total extracted mass as a first approximation to gauge reproducibility. The concentration of coffee is phenomenologically related to the refractive index of the beverage (which is temperature dependent), and can be recovered using a previously presented methodology.[19]

We elected to use a representative modern espresso recipe (*i.e.* 20.0(5) g of dry ground coffee in, 40.0(5) g of beverage out). Temperature was kept constant, 92 °C. Espresso shots were discarded if, due to human error, the shot mass was outside a ±1 g tolerance. Exact beverage masses were included in each calculation of extraction yield. Calibrating measurements were made using a larger sample size, $n = 20$ and subsequent data was collected in pentaplicate.

Laser diffraction particle size analysis was performed on a Beckman Coulter LS13 320 MW. The instrument has a built-in dark field reticule which is used to ensure correct optical alignment. An alignment check was carried out prior to every run to ensure the optimum accuracy of the particle size distribution.

Nitrogen physisorption isotherm data was acquired at -196 C on a Quadrasorb SI (Quantachrome Instruments). Prior to measurement, each ground coffee sample was degassed twice at 200 C under vacuum for 16 hours. As a consequence of the minimal gas adsorption (due to low surface area and a lack of small pores) at partial pressures < 0.3, it was not possible to obtain a linear fit to the Brunaueur–Emmett–Teller (BET) equation. Therefore, a specific BET surface area value is not reported.

X. AUTHOR CONTRIBUTIONS

The study was conceived by C.H.H., M.I.C., J.M.F, and W.T.L. M.I.C. and D.M. brewed the espresso. D.H. and E.U. performed the grinding experiments. Z.C.K and S.A.F. performed the gas sorption measurements. J.W., W.T.L, and J.M.F. formulated the mathematical model. C.H.H., W.T.L., and J.M.F. wrote the manuscript and all authors contributed to the final version.

[1] W. B. Sunarharum, D. J. Williams, and H. E. Smyth, "Complexity of coffee flavor: A compositional and sensory perspective," Food Res. Int. **62**, 315–325 (2014).

[2] M. Spencer, E. Sage, M. Velez, and J.-X. Guinard, "Using single free sorting and multivariate exploratory methods to design a new coffee taster's flavor wheel: design of coffee taster's flavor wheel," J. Food Sci. **81**, S2997–S3005 (2016).

[3] S. Schenker, C. Heinemann, M. Huber, R. Pompizzi, R. Perren, and R. Escher, "Impact of roasting conditions on the formation of aroma compounds in coffee beans," J. Food Sci. **67**, 60–66 (2002).

[4] J. Baggenstoss, L. Poisson, R. Kaegi, R. Perren, and F. Escher, "Coffee roasting and aroma formation: application of different time- temperature conditions," J. Agric. Food Chem. **56**, 5836–5846 (2008).

[5] S. Rao, *The Roaster's Companion* (Self Published, 2014).

[6] S. Rao, *The Professional Barista's Handbook* (Self Published, 2008).

[7] S. Rao, *Everything But Espresso: Professional Coffee Brewing Techniques* (Self Published, 2010).

[8] National Coffee Association of USA, "Understanding the economic impact of the united states coffee industry," (Accessed 07 December 2019),

http://www.ncausa.org/Industry-Resources/Economic-Impact/Economic-Impact-Infographic.
[9] C. Gay, F. Estrada, C. Conde, H. Eakin, and L. Villers, "Potential impacts of climate change on agriculture: a case of study of coffee production in veracruz, mexico," Clim. Change **79**, 259–288 (2006).
[10] J. Jaramillo, E. Muchugu, F. E. Vega, A. Davis, C. Borgemeister, and A. Chabi-Olaye, "Some like it hot: the influence and implications of climate change on coffee berry borer (hypothenemus hampei) and coffee production in east africa," PLoS One **6**, e24528 (2011).
[11] A. P. Davis, T. W. Gole, S. Baena, and J. Moat, "The impact of climate change on indigenous arabica coffee (coffea arabica): predicting future trends and identifying priorities," PloS One **7**, e47981 (2012).
[12] C. Bunn, P. Läderach, O. O. Rivera, and D. Kirschke, "A bitter cup: climate change profile of global production of arabica and robusta coffee," Clim. Change **129**, 89–101 (2015).
[13] E. Uman, M. Colonna-Dashwood, L. Colonna-Dashwood, M. Perger, C. Klatt, S. Leighton, B. Miller, K. T. Butler, B. C. Melot, R. W. Speirs, and C. H. Hendon, "The effect of bean origin and temperature on grinding roasted coffee," Sci. Rep. **6**, 24483 (2016).
[14] A. N. Glöss, B. Schönbächler, M. Rast, L. Deuber, and C. Yeretzian, "Freshness indices of roasted coffee: Monitoring the loss of freshness for single serve capsules and roasted whole beans in different packaging," CHIMIA International Journal for Chemistry **68**, 179–182 (2014).
[15] C. F. Ross, K. Pecka, and K. Weller, "Effect of storage conditions on the sensory quality of ground arabica coffee," J. Food Qual. **29**, 596–606 (2006).
[16] A. Farah, "Coffee constituents," in *Coffee*, edited by Y.-F. Chu (John Wiley & Sons, Oxford, UK, 2012) pp. 21–58.
[17] B. Folmer, *The craft and science of coffee* (Academic Press, 2016).
[18] Chan-Yuan Tan and Yao-Xiong Huang, "Dependence of refractive index on concentration and temperature in electrolyte solution, polar solution, nonpolar solution, and protein solution," J. Chem. Eng. Data **60**, 2827–2833 (2015).
[19] V. Fedele, "Universal refractometer apparatus and method," (2012), US Patent 8,239,144.
[20] P. Kanouté, D. P. Boso, J. L. Chaboche, and B. A. Schrefler, "Multiscale methods for composites: a review," Arch. Comput. Methods Eng. **16**, 31–75 (2009).
[21] E-Imports, "Coffee statistics," (Accessed 07 December 2019), http://www.e-importz.com/coffee-statistics.php.
[22] M. Spiro and R. M. Selwood, "The kinetics and mechanism of caffeine infusion from coffee: the effect of particle size," J. Sci. Food Agric. **35**, 915–924 (1984).
[23] M. Spiro and C. M. Page, "The kinetics and mechanism of caffeine infusion from coffee: hydrodynamic aspects," J. Sci. Food Agric. **35**, 925–930 (1984).
[24] M. Spiro and J. E. Hunter, "The kinetics and mechanism of caffeine infusion from coffee: the effect of roasting," J. Sci. Food Agric. **36**, 871–876 (1985).
[25] M. Spiro, R. Toumi, and M. Kandiah, "The kinetics and mechanism of caffeine infusion from coffee: the hindrance factor in intra-bean diffusion," J. Sci. Food Agric. **46**, 349–356 (1989).
[26] M. Spiro, "Modelling the aqueous extraction of soluble substances from ground roast coffee," J. Sci. Food Agric. **61**, 371–372 (1993).
[27] M. Spiro and Y. Y. Chong, "The kinetics and mechanism of caffeine infusion from coffee: the temperature variation of the hindrance factor," J. Sci. Food Agric. **74**, 416–420 (1997).
[28] B. R. Corrochano, J. R. Melrose, A. C. Bentley, P. J. Fryer, and S. Bakalis, "A new methodology to estimate the steady-state permeability of roast and ground coffee in packed beds," J. Food Eng. **150**, 106 – 116 (2015).
[29] "Coffee beans," (Accessed 07 December 2019), www.optics.rochester.edu/workgroups/cml/opt307/spr16/.
[30] K. M. Moroney, *Heat and mass transfer in dispersed two-phase flows*, Ph.D. thesis, University of Limerick (2016).
[31] A. C. Fowler and B. Scheu, "A theoretical explanation of grain size distributions in explosive rock fragmentation," Proc. R. Soc. A **472**, 20150843 (2016).
[32] J. M. Foster, S. J. Chapman, G. Richardson, and B. Protas, "A mathematical model for mechanically-induced deterioration of the binder in lithium-ion electrodes," SIAM J. Appl. Math. **77**, 2172–2198 (2017).
[33] G. Richardson, G. Denuault, and C. P. Please, "Multiscale modelling and analysis of lithium-ion battery charge and discharge," J. Eng. Math. **72**, 41–72 (2012).
[34] J. M. Foster, A. Gully, H. Liu, S. Krachkovskiy, Y. Wu, S. B. Schougaard, M. Jiang, G. Goward, G. A. Botton, and B. Protas, "Homogenization study of the effects of cycling on the electronic conductivity of commercial lithium-ion battery cathodes," J. Phys. Chem. C **119**, 12199–12208 (2015).
[35] D. A. G. von Bruggeman, "Berechnung verschiedener physikalischer konstanten von heterogenen substanzen. i. dielektrizitätskonstanten und leitfähigkeiten der mischkörper aus isotropen substanzen," Ann. physik **416**, 636–664 (1935).
[36] J. O. Robinson, "The misuse of taste names by untrained observers," British J. Psych. **61**, 375–378 (1970).
[37] A. Cruz and B. G. Green, "Thermal stimulation of taste," Nature **403**, 889–892 (2000).
[38] R. A. Clará, A. C. G. Marigliano, and H. N. Sólimo, "Density, viscosity, and refractive index in the range 283.15 to 353.15 k and vapor pressure of α-pinene, d-limonene, $\pm$-linalool, and citral over the pressure range 1.0 kpa atmospheric pressure," J. Chem. & Eng. Data **54**, 1087–1090 (2009).
[39] G. Allaire, "Homogenization of the stokes flow in a connected porous medium," Asymp. Anal. **2**, 203–222 (1989).
[40] D. Polisevsky, "On the homogenization of fluid flows through porous media," Rend. Sem. Mat. Univ. Politecn. Torino **44**, 383–393 (1986).
[41] N. Dukhan, Ö. Bağcı, and M. Özdemir, "Experimental flow in various porous media and reconciliation of forchheimer and ergun relations," Exp. Therm. Fluid Sci. **57**, 425–433 (2014).
[42] A. C. Fowler, B. Scheu, W. T. Lee, and M. J. McGuinness, "A theoretical model of the explosive fragmentation of vesicular magma," Proc. Royal Soc. Lon. A **466**, 731–752 (2010).
[43] J. Kestin, M. Sokolov, and W. A. Wakeham, "Viscosity of liquid water in the range - 8 c to 150 c," J. Phys. Chem. Ref. Data **7**, 941–948 (1978).
[44] A. S. Franca, L. S. Oliveira, J. C. F. Mendonça, and X. A. Silva, "Physical and chemical attributes of defective crude and roasted coffee beans," Food Chem. **90**, 89–94 (2005).

[45] Y. Zeng, P. Albertus, R. Klein, N. Chaturvedi, A. Kojic, M. Z. Bazant, and J. Christensen, "Efficient conservative numerical schemes for 1d nonlinear spherical diffusion equations with applications in battery modeling," J. Electrochem. Soc. **160**, A1565–A1571 (2013).
[46] J. M. Foster, "Espresso simulation code," `https://github.com/jamiemfoster/Espresso` (2019), accessed 07 December 2019.

Appendix A: Multiple scales homogenization

In this section we carry out upscaling (or homogenization) of the system of equations formulated in the Supplemental Information. Rather than resorting to a lengthy multiple scales analysis we use the previous results from the rigorous analysis[33] to formulate the multi-scale system of equations. Before proceeding we note the abuse of notation used in this section and henceforth; namely, that the dependent variables here, and in §II B are not strictly the same as those appearing in §II but are instead their homogenized counterparts. Despite this, in the interests of brevity we opt not to embed this distinction within the notation.

Upscaling the Navier-Stokes equations at suitably small Reynolds numbers (so that flow is laminar) on a porous media results in Darcy's law[39,40], whilst upscaling (18) and accounting for the source/sink of solutions into/out of the liquid arising from the boundary conditions (14) results in a reaction-advection-diffusion equation. The calculation can be adapted appropriately[33] to show that

$$\mathbf{q} = -\frac{\kappa}{\mu}\nabla P, \quad \text{(A1)}$$

$$(1-\phi_s)\frac{\partial c_l^*}{\partial t} = \nabla\cdot(D_{eff}\nabla c_l^* - \mathbf{q}c_l^*) + b_{et,1}G_1 + b_{et,2}G_2, \quad \text{(A2)}$$

where the star has been appended to c_l to emphasize that it has undergone the homogenization procedure. Here, κ is the permeability of the packed bed and $\mathbf{q}$ is the Darcy flux (*i.e.*, the discharge per unit area with units of m/s) which is related to the average fluid velocity within the pore space, ν, via $q = (1-\phi_s)\nu$. One might conjecture that in espresso making applications pressures are sufficiently high, and the pores sufficiently small, that turbulent flow could be present. However, the experimental evidence strongly indicates that a model based on Darcy's law (rather than Ergun or Forchheimer[41,42]) is suitable[28]. In (A2), $b_{et,i}$ are the Brunauer-Emmett-Teller surface areas (defined to be the surface area, of grain species i, per unit volume of puck), and D_{eff} is the effective macroscopic diffusivity. The quantity ϕ_s appearing in (A2) is the local volume fraction of coffee grains which, under our assumption that the grains are spheres, is related to the BET surface areas and radii by

$$\phi_s = \phi_{s,1} + \phi_{s,2}, \quad \text{where} \quad \phi_{s,i} = \frac{b_{et,i}}{4\pi a_i^2}\frac{4}{3}\pi a_i^3. \quad \text{(A3)}$$

The latter equation is a product of: (i) the surface area of particles per volume of bed divided by the surface area of a single particle—which is therefore the number of particles of a particular type per volume of bed, and; (ii) the volume of a single particle of a particular type. The product of these two factors is therefore the local volume fraction of the bed of particles of a particular size, and the sum over the terms gives the local solid volume fraction ϕ_s.

The form of the boundary and initial conditions to close (A2) remain unchanged, and are (7), (10) and (11) with the exception that $\mathbf{u}$ is interchanged with $\mathbf{q}$. An important consequence of upscaling the Navier-Stokes equations to give Darcy's law is that the order of the system is reduced. We are therefore not at liberty impose all the boundary conditions laid out in §II. To close (A1) we retain (5), (8) and the component of (12) normal to the boundary, *i.e.*, $\mathbf{q}\cdot\hat{\mathbf{n}}|_{R=R_0} = 0$. The omission of the remaining conditions is justified on the basis that, had we carried out the homogenization explicitly, we would have imposed it when solving on the microscopic length scale.

The governing equations for the coffee concentration in the solid phase remain unchanged by the homogenization procedure; c_{si} are still to be determined by solving (4) with their associated boundary and initial conditions, (13) and the latter equation in (17). In the context of the multiple scales approach one can think of the retention of the full system for c_{si} as being necessitated by the need to evaluate the reaction rates, G_i, which can only be done by returning to the "microscopic" scale. One must therefore solve two equations of the form (4) at each station in macroscopic coordinate system, one with $i = 1$ the other with $i = 2$, to predict the coffee concentration profiles within a representative coffee grain of radius a_i (for $i = 1, 2$).

Appendix B: One-dimensional reduction

It follows from the form of the upscaled equations and boundary conditions that the *marcoscopic* solution should be one-dimensional, *i.e.*, that it should depend only on z, depth through the basket. In this case Darcy's equation reduces to

$$q = -\frac{\kappa}{\mu}\frac{\partial P}{\partial z}, \quad \text{(B1)}$$

where the scalar q is the z-component of $\mathbf{q}$.

Since the primary focus of the model is to capture the extraction process after the initial wetting phase we assume that in the regime of interest the flow is steady. Despite this simplification we will provide scope for the model to capture consolidation of the bed that may occur in the wetting phase where the finer grains may be swept towards the bottom of the basket by the intruding liquid — in the coffee industry this process is known as fines migration. Importantly for our model such a phenomena

would lead to spatial variations in the bed permeability. The solution to (B1), with spatial dependence in κ, supplemented by the boundary conditions (5) and (8) is

$$q = \frac{\kappa_{eff} P_{tot}}{\mu L}, \qquad \kappa_{eff} = \frac{1}{L}\int_0^L \kappa(s) ds. \tag{B2}$$

Inserting the result (B2) into (A1) and reducing to dependence only on z gives equations the system (18)-(19) for c_l^*. On assuming radial symmetry within each coffee grain the resulting problems to be solved for c_{s1} and c_{s2} are (21)-(22). The problem is closed by supplying the initial conditions (17).

Appendix C: Tuning the model to espresso extraction data

Table S1 summarizes the parameter values harvested from the literature. The values shown in Table S2 were derived by choosing a demarcation between fines and boulders of 100 μm and integrating the grind data shown in Figure 1 leading to the volume contributions of the two families of particles. This was then converted into BET surface area contributions using the relationship (A3). Intermediate grind settings were linearly interpolated from these data.

Appendix D: Numerical approach

Here, an approach to numerically solve (17)-(22) is described. The method centres around: (i) finite difference approximations of spatial derivatives in z; (ii) the use of a control volume method[45] for treatment of the spatial dependence in r; (iii) and, MATLAB's ODE suite for temporal integration. The code is hosted on a GitHub repository and is available.[46]

Finite differences are the method of choice for the derivatives in z primarily for their ease of implementation. The particular control volume method chosen is particularly apt to treat the transport of coffee in the solid phase for two reasons. Firstly, because it exhibits perfect conservation of coffee mass, which can be hard to ensure using standard finite differences owing to the singularity at the origin of the radial coordinate in the spherical diffusion equation. Secondly, in contrast to many other control volume methods, it provides direct access to the concentration on the surface of the particles thereby avoiding the need for extrapolation which would inevitably introduce additional errors. This surface concentration determines the reaction rate across the solid grain boundaries and so accurate evaluation of this quantity is crucial for reliable simulation. After applying these treatments for the spatial dependencies, the system of PDEs (17)-(22) are reduced to system of coupled ODEs. We select the MATLAB routine `ode15s` to integrate this system forward in time because: (i) it is able to cope with solving a system of differential-algebraic equations; (ii) it offers adaptive time-stepping, and; (iii) has relatively modest computational cost.

We introduce N equally spaced grid points, z_j for $j \in [1, N]$, thereby dividing the spatial z-domain into $N-1$ equally spaced subdomains. The grid spacing in z is therefore given by $h_z = 1/(N-1)$. Henceforth we adopt the shorthand notation $c_l^*(z,t)|_{z=z_j} = c_{l,j}^*(t)$. At each station in z we must solve for the coffee concentration within a representative grain, *i.e.* , at each z_j we must solve two equations of the form (21); one with $i = 1$ for c_{s1} and another with $i = 2$ for c_{s2}. Each of the $2N$ copies of equation (21) are discretized by introducing M equally spaced grid points, r_k for $k \in [1, M]$, which subdivide each r-domain into $M-1$ subdomains. The grid spacing in r is therefore given by $h_r = 1/(M-1)$. In total we have $N \times M$ different stations in r and at these locations we denote the value of the coffee concentration in small and large particles using the following shorthands $c_{s1}(z,r,t)|_{z=z_j, r=r_k} = c_{s1,j,k}(t)$ and $c_{s2}(z,r,t)|_{z=z_j, r=r_k} = c_{s2,j,k}(t)$, respectively. Thus, the index j indicates the representative particle's position in z whereas k labels the radial position within a particular representative particle. The $(2M+1) \times N$ unknown functions of time were converted into one large column vector $\mathbf{u}(t)$ as follows

$$\begin{aligned}\mathbf{u}(t) = [\mathbf{c}_l(t)^{*\mathrm{T}}\ \mathbf{c}_{s1,1}(t)^{\mathrm{T}}\ \mathbf{c}_{s1,2}(t)^{\mathrm{T}} \\ \cdots\ \mathbf{c}_{s1,N}(t)^{\mathrm{T}}\ \mathbf{c}_{s2,1}(t)^{\mathrm{T}}\ \mathbf{c}_{s2,2}(t)^{\mathrm{T}}\ \cdots\ \mathbf{c}_{s2,N}(t)^{\mathrm{T}}]^{\mathrm{T}}.\end{aligned} \tag{D1}$$

We now rewrite the problem in the form

$$\mathbf{M}\frac{d\mathbf{u}}{dt} = \mathbf{f}(\mathbf{u}), \tag{D2}$$

where $\mathbf{M}$ is the mass matrix and $\mathbf{f}(\mathbf{u})$ is a nonlinear function which arises from the application of finite difference approximations (to the equations in z) and the control volume method (for the equations in r). The system of ODEs (D2) is written in the standard form accepted by MATLABs `ode15s`.

Below we present the first N entries of the nonlinear function $\mathbf{f}(\mathbf{u})$ arising from the discretisation of (18) and its boundary conditions (19). The remaining $2MN$ equations arising from applying control volumes to (21) are previously detailed[45], and so in the interests of brevity

Description	**Sym.**	**Val.**	**Ref.**
Radius of bed	R_0	29.2×10^{-3}m	—
Viscosity of water at 90°C	μ	3.15×10^{-4} Pa·s	[43]
Density of water 90°C	ρ_{out}	997 kg/m^3	—
Roasted coffee bulk density	$\rho_{grounds}$	330 kg/m^3	[44]
Saturation concentration of water	c_{sat}	212.4 kg/m^3	[30]
Initial concentration of solubles in grounds	c_{s0}	118.0 kg/m^3	[30]
Volume fraction of grounds in packed bed	ϕ_s	0.8272	[30]
Dose of grounds "in"	M_{in}	20 g	—
Mass of beverage "out"	M_{out}	40 g	—
Pump overpressure	P_{tot}	5 bar	—

TABLE S1. Parameter values above the horizontal score were taken from the literature, and those below were chosen to mimic the experimental extraction protocol used here.

Grind setting, G_s	1.0	1.5	2.0	2.5
Vol. fraction of fines, ϕ_{s1}	0.1689	0.1343	0.1200	0.0780
Vol. fraction of boulders, ϕ_{s1}	0.6583	0.6929	0.7072	0.7492
Radius of boulders, a_2 **(μm)**	273.86	335.41	335.41	410.79
BET surface area of fines, b_{et1} **(1/m)**	10.187×10^3	8.099×10^3	7.239×10^3	4.703×10^3
BET surface area of boulders, b_{et2} **(1/m)**	7.211×10^3	6.197×10^3	6.325×10^3	5.472×10^3

TABLE S2. Parameter values above the horizontal score were extracted directly from the experimental data shown in Figure 2 and those below were subsequently inferred using (A3). Throughout, we take the radius of fines to be fixed with $a_1 = 12\mu$m.

we do not repeat them here. We have

$$M_{1,1} = 0, \quad \text{(D3)}$$

$$f_1 = -\frac{D_{eff}}{h_z}\left(-\frac{3}{2}u_1 + 2u_2 - \frac{1}{2}u_3\right) + u_1, \quad \text{(D4)}$$

$$M_{i,i} = 1 - \phi_s, \quad \text{(D5)}$$

$$f_i = \frac{D_{eff}}{h_z^2}\left(u_{i+1} - 2u_i + u_{i-1}\right) - \frac{u_{i+1} - u_{i-1}}{2h_z} + b_{et,1}K, \quad \text{(D6)}$$

$$M_{N,N} = 0, \quad \text{(D7)}$$

$$f_N = -\frac{D_{eff}}{h_z}\left(\frac{1}{2}u_{N-2} - 2u_{N-1} + \frac{3}{2}u_N\right). \quad \text{(D8)}$$

Appendix E: Supplemental Experiments

The BET sorption data and effect of tamp force are presented.

Grind setting, G_s	1.1	1.3	1.5	1.7	1.9	2.1	2.3
Shot time, t_{shot} **(s)**	37.6	35.7	33.9	30.5	28.5	26.6	24.0
Darcy flux, q **(m/s)**	3.98×10^{-4}	4.20×10^{-4}	4.42×10^{-4}	4.91×10^{-4}	5.26×10^{-4}	5.63×10^{-4}	6.24×10^{-4}

TABLE S3. Parameter values above the horizontal score are experimental values shown in Figure 4 whilst the values of the Darcy flux below the horizontal score were computed via (26).

Dose of grounds "in", M_{in} **(g)**	16	18	20	22	24
Bed depth, L **(m)**	15.0×10^{-3}	16.8×10^{-3}	18.7×10^{-3}	20.6×10^{-3}	22.5×10^{-3}

TABLE S4. Parameter values that were adjusted to explore the effects of altering the dose of grounds "in".

	Grind setting, G_s	1.1	1.3	1.5	1.7	1.9	2.1	2.3
Pump overpressure, $P_{tot} = 3$ **bar**	**Darcy flux,** q **(m/s)**	2.39×10^{-4}	2.53×10^{-4}	2.65×10^{-4}	2.95×10^{-4}	3.16×10^{-4}	3.38×10^{-4}	3.74×10^{-4}
Pump overpressure, $P_{tot} = 5$ **bar**	**Darcy flux,** q **(m/s)**	3.98×10^{-4}	4.20×10^{-4}	4.42×10^{-4}	4.91×10^{-4}	5.26×10^{-4}	5.63×10^{-4}	6.24×10^{-4}
Pump overpressure, $P_{tot} = 7$ **bar**	**Darcy flux,** q **(m/s)**	5.57×10^{-4}	5.88×10^{-4}	6.19×10^{-4}	6.87×10^{-4}	7.36×10^{-4}	7.88×10^{-4}	8.74×10^{-4}
Pump overpressure, $P_{tot} = 9$ **bar**	**Darcy flux,** q **(m/s)**	7.16×10^{-4}	7.56×10^{-4}	7.96×10^{-4}	8.84×10^{-4}	9.47×10^{-4}	10.13×10^{-4}	11.23×10^{-4}

TABLE S5. Parameter values that were adjusted to explore the role of altering the pump overpressure. The values of the Darcy flux for a 5 bar overpressure are identical to those on Table S3 and the others have been computed from those via the relationship (26).

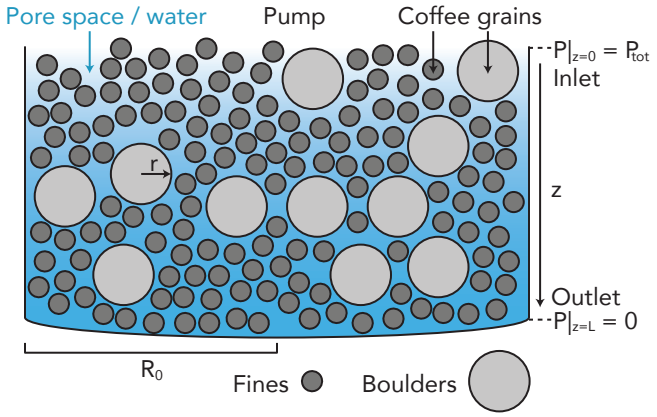

FIG. 1. A schematic of the espresso basket model geometry showing the coffee grounds in grey (Ω_s), and the pore space which is filled with water during extraction in blue (Ω_l). The macroscopic spatial coordinate measuring depth through the bed, z, the microscopic spatial coordinate measuring radial position within the spherical coffee particles, r, as well as the basket radius, R_0, are also indicated.

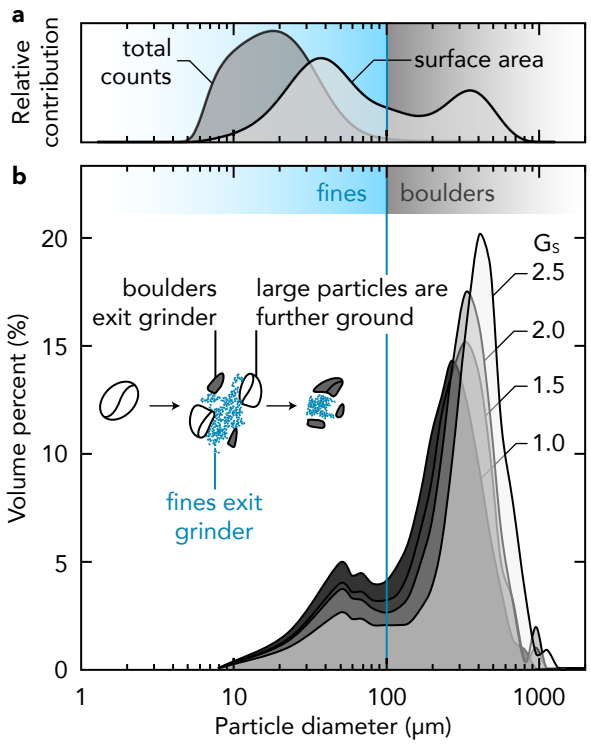

FIG. 2. Particle size distributions collected using the method described in our experimental procedure. (**a**) Surface area and number of coffee particulates produced with a grind setting $G_S = 2.5$. Here, 99% of the particles are <100 μm in diameter, and account for 80% of the surface area. (**b**) The volume percent particle size distribution at $G_S = 2.5$, 2.0, 1.5, and 1.0. Grinding finer reduces the average boulder size, while increasing the number of fines. Intruders are boulders that are larger than the aperture of the burr set and hence further fractured until they can exit the burrs.

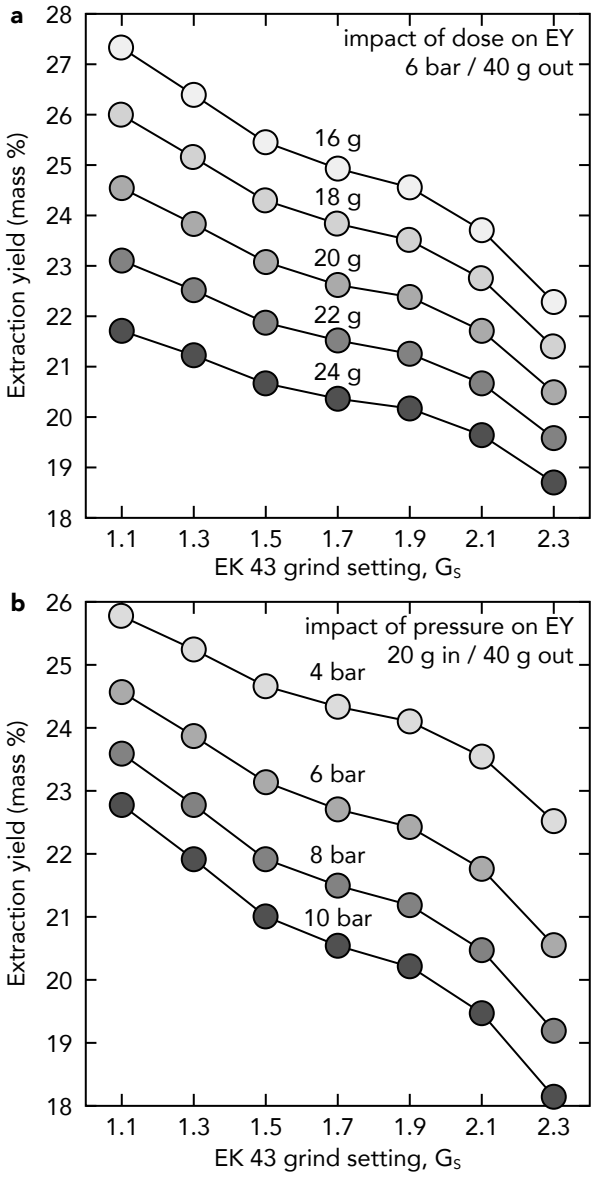

FIG. 3. Extraction yield as a function of grind size, with varying coffee dose and water pressure. (**a**) The effect of changing coffee dose M_{in} with constant water pressure shows that reducing initial coffee mass, but keeping the beverage volume constant, results in higher extractions. (**b**) The effect of changing pump overpressure, P, with a constant brew ratio shows an increase in extraction yield with decrease in water pressure.

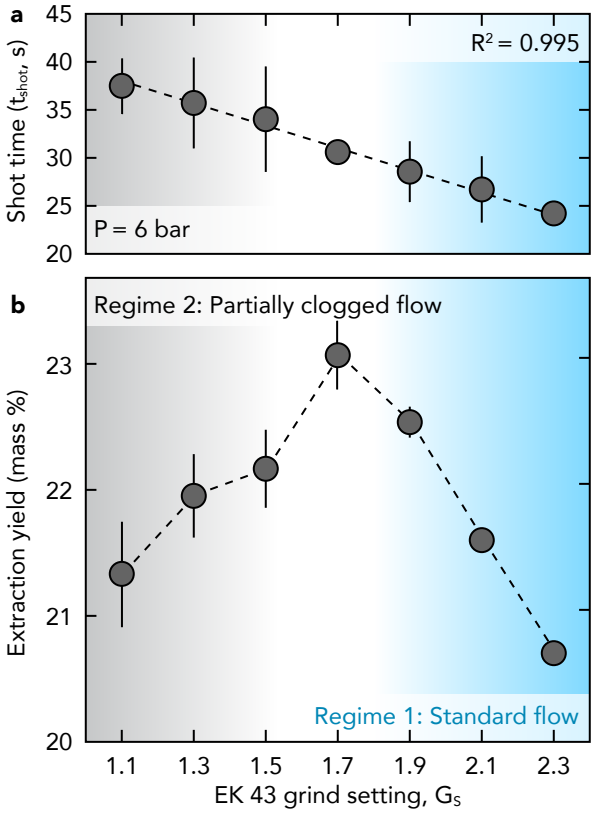

FIG. 4. (**a**) $P_W = 6$ bar, $t_F = 98$ N shot times are inversely proportional to G_S. (**b**) Extracted mass percent can be described by two regimes. Regime 1: a standard flow system where an expected increase in extraction percent is observed with reducing G_S. Conversely, Regime 2: partially clogged flow is operative when there are too many fines (*ca.* $G_S = 1.7$), forming aggregates/inhomogeneous bed density, effectively reducing the surface area of the granular bed.

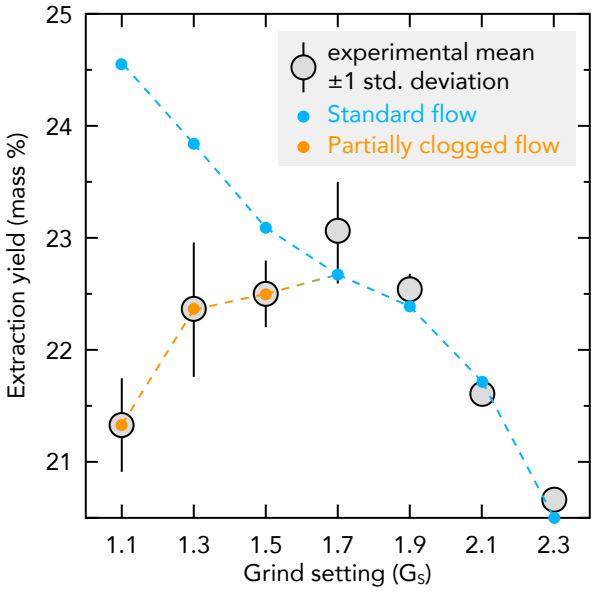

FIG. 5. A comparison of the model and experimental espresso shots collected using a standard recipe of 20 g coffee dry mass, 40 g beverage mass, produced using 6 bar water pressure. The partially clogged and standard flow regimes are highlighted in orange and blue, respectively.

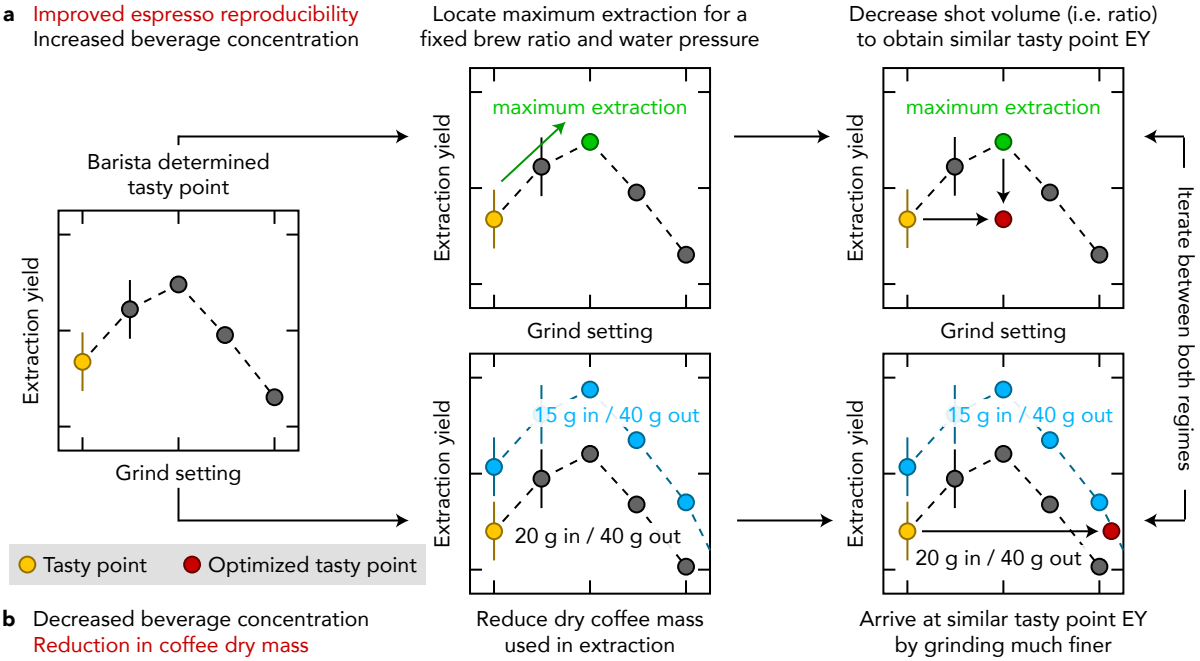

FIG. 6. Information gleaned from the model provide two strategies to improve espresso reproducibility. After determining a tasty point (yellow), the barista can obtain the same EY by, **a**) grinding coarser to find the maximum EY (green) and then reducing water mass, or **b**) down-dosing and grinding much coarser. The former results in smaller beverages with higher coffee concentration, while the latter results in less dry mass coffee being used, at lower concentration. Here "in" and "out" refer to dry coffee mass and beverage volume, respectively.

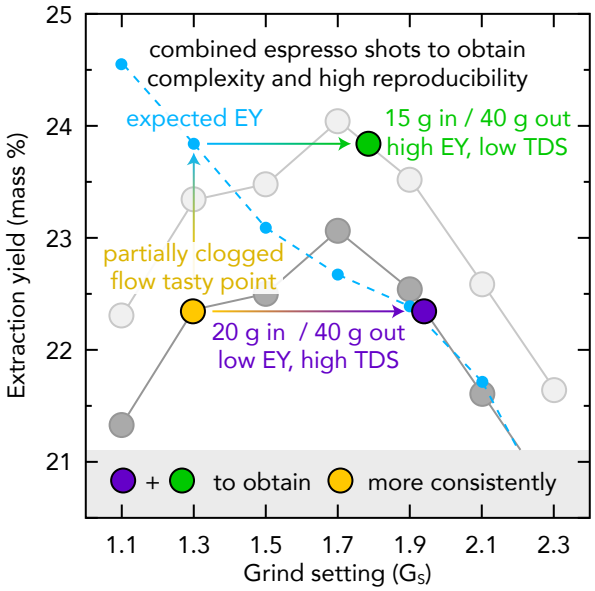

FIG. 7. Blending two espresso shots provides an avenue to obtaining the flavor complexity of the partially clogged flow tasty point (yellow) with less shot variation. One high dry mass, low EY shot (purple) combined with one low dry mass, high EY shot (green) provide an approximation to the tasty point.